

# Discussion Paper

Central Bureau of Statistics, P.B. 8131 Dep, 0033 Oslo 1, Norway

No. 12

21 June 1985

## *MARKOV CHAINS GENERATED BY MAXIMIZING COMPONENTS OF MULTIDIMENSIONAL EXTREMAL PROCESSES*

BY

*JOHN K. DAGSVIK*

### ABSTRACT

A multidimensional inhomogenous extremal process is defined and it is demonstrated that it belongs to the class of pure jump Markov processes. Let  $\{Z_j(t)\}$  be the  $j$ -th component of the process. Let  $\{J(t)\}$  be a finite state process defined by  $J(t) = j$  if  $Z_j(t) = \max_k Z_k(t)$ . It is proved that  $\{J(t)\}$  is an inhomogenous Markov chain and the transition probabilities of this chain are obtained. The chain  $\{J(t)\}$  provides a framework for modelling mobility processes that are generated from intertemporal utility-maximizing individuals.

Not to be quoted without permission from author(s). Comments welcome.

## 1. Introduction

The multidimensional extremal process has been defined and examined by de Haan and Resnick [3]<sup>1)</sup>. The study of extremal processes can be motivated as follows:

Let  $\tilde{X}_i = (\tilde{X}_{i1}, \tilde{X}_{i2}, \dots, \tilde{X}_{im})$ ,  $i = 1, 2, \dots$ , be a sequence of independent identically distributed random vectors. Define the processes  $U_{nj}(t) = \max_{i \leq [nt]} \tilde{X}_{ij}$ . Suppose there exists constants  $a_{nj}$  and  $b_{nj}$  such that

$$\tilde{Z}_n(t) = \left( \frac{U_{n1}(t) - b_{n1}}{a_{n1}}, \frac{U_{n2}(t) - b_{n2}}{a_{n2}}, \dots, \frac{U_{nm}(t) - b_{nm}}{a_{nm}} \right)$$

converges weakly to a stochastic process  $\{\tilde{Z}(t)\}$ . Then  $\{\tilde{Z}(t)\}$  belongs to the class of multidimensional extremal processes.

Consider a multidimensional extremal process,  $\{\tilde{Z}(t)\} = \{Z_1(t), Z_2(t), \dots, Z_m(t)\}$ . Define a finite state space process  $\{J(t)\}$  where  $J(t) = j$  if  $Z_j(t) = \max_k Z_k(t)$ . In the case when  $\{Z_k(t)\}$ ,  $k = 1, 2, \dots, m$ , are independent extremal processes, it is shown in [1] that  $\{J(t)\}$  is a Markov chain. As a consequence, the difference between two independent extremal processes has exponentially distributed excursion times because they are the holding times of  $\{J(t)\}$ . (Recall that excursion times are the time intervals the process lies below or above a given level.)

The process  $\{J(t)\}$  is of substantial interest in a variety of applications in psychology and economics. Consider the following motivating example. Each individual of a population has the choice between different careers. At each point in time the individuals have the choice between  $j = 1, 2, \dots, n$ , alternatives (states). Assume that the

1) Their definition differ from the multivariate extremal process studied by Weissman [10].

attractiveness of state  $j$  is measured by a latent index  $Z_j(t)$  (utility) at time  $t$ . The individual decision rule is to move to the state with the highest utility at that time. The utility process  $\{Z_j(t)\}$ , is considered random because not all the variables that influence the individuals' choice are observable to the observer. From the observer's point of view the decision process is exactly the process  $\{J(t)\}$ .

At any given point in time the probability of being in a particular state takes the multinomial logit form provided  $Z_1(t), Z_2(t), \dots$  are independent. Since the logit model is consistent with a famous axiom from mathematical psychology called "independence from irrelevant alternatives" (IIA) (cf. [6]) it provides a behavioral justification for independent extreme value distributed utilities. However, in many applications it may be implausible to require the IIA property to hold. This has lead to the development of choice models generated from general extreme value distributed utilities, (see [7]).

In Dagsvik [1] the process  $\{J(t)\}$  was studied in the case where the components of  $\{Z(t)\}$  are independent processes. The purpose of the present paper is to extend these results to allow for interdependent components,  $Z_j(t)$ ,  $j=1,2,\dots,m$ .

## 2. Preliminaries

Let  $\{F_t, t \geq 0\}$  be a family of multidimensional extreme value distribution functions that satisfies  $F_0=1$  and

$$(2.1) \quad G_t(x_1, x_2, \dots, x_m) = e^{-y} G_t(x_1 - y, x_2 - y, \dots, x_m - y), \quad \forall y,$$

where  $G_t = -\log F_t$ . Condition (2.1) implies that the univariate marginals

have the form  $\exp\{-Ce^{-x}\}$  which is the type III extreme value distribution (see Johnson and Kotz [5]). Conditions that allow for type I and II marginals will be considered in Section 3.

Suppose furthermore that  $F_t/F_s$  is a nondecreasing function in  $(x_1, x_2, \dots, x_m)$  for  $s < t$ .

Let  $\{\tilde{W}(s, t)\}$ ,  $0 \leq s \leq t$ , be a family of  $m$ -dimensional vector variables with law

$$(2.2) \quad P(W_1(s, t) \leq x_1, W_2(s, t) \leq x_2, \dots, W_m(s, t) \leq x_m) \\ = F_t(x_1, x_2, \dots, x_m) / F_s(x_1, x_2, \dots, x_m)$$

and with the property that when  $(s, t) \cap (s', t') = \emptyset$  then  $\tilde{W}(s, t)$  and  $\tilde{W}(s', t')$  are independent. Let  $0 = t_0 < t_1 < t_2 < \dots < t_n$  be arbitrary points in time. Define a stochastic process  $\{\tilde{Z}(t), t \geq 0\}$  recursively by

$$(2.3) \quad \tilde{Z}(t) = \max(\tilde{Z}(s), \tilde{W}(s, t)), \quad s < t, \quad \tilde{Z}(0) = -\infty$$

where maximum is taken componentwise. From (2.2) and (2.3) we obtain the finite dimensional marginal distribution of  $\{\tilde{Z}(t)\}$ :

$$(2.4) \quad P\left\{\bigcap_{j=1}^n (Z_1(t_j) \leq x_j(1), Z_2(t_j) \leq x_j(2), \dots, Z_m(t_j) \leq x_j(m))\right\} \\ = \prod_{j=1}^n F_{t_j}(u_j(1), u_j(2), \dots, u_j(m)) / F_{t_{j-1}}(u_j(1), u_j(2), \dots, u_j(m))$$

where

$$u_j(k) = \min_{i \geq j} x_i(k), \quad k=1, 2, \dots, m, \quad j=1, 2, \dots, n.$$

It is immediately seen from (2.3) that  $\{\tilde{Z}(t)\}$  is nondecreasing.

We call  $\{\tilde{Z}(t)\}$  a multidimensional inhomogeneous extremal process. The definition presented here is a direct extension of the one-dimensional case discussed by Weissman [9].

An immediate consequence of (2.3) is that  $\{Z(t)\}$  is a Markov process. However,  $\{Z(t)\}$  also possesses a particular "extended" Markov property stated below. For simplicity it is stated only for  $m=2$ .

Lemma 1: Let  $x_1(k) \leq x_2(k) \leq \dots \leq x_n(k)$ ,  $k=1,2$  and let  $B_j(k)$  denote  $\{x_j(k)\}$  or  $(-\infty, x_j(k)]$  for  $j=1,2,\dots,n-1$ , and  $k=1,2$ . We have

$$(2.5) \quad P\{Z_1(t_n) \leq x_n(1), Z_2(t_n) \leq x_n(2) \mid Z_1(t_j) \in B_j(1), Z_2(t_j) \in B_j(2), j=1,2,\dots,n-1\} \\ = P\{Z_1(t_n) \leq x_n(1), Z_2(t_n) \leq x_n(2) \mid Z_1(t_{n-1}) \in B_{n-1}(1), Z_2(t_{n-1}) \in B_{n-1}(2)\}.$$

Proof: If  $x_j(k) \leq x_{j+1}(k)$  for  $j=1,2,\dots$ , and  $k=1,2$ , we get from (2.4)

$$P\left\{\bigcap_{j=1}^n (Z_1(t_j) \leq x_j(1), Z_2(t_j) \leq x_j(2))\right\} = \prod_{j=1}^n F_{t_j}(x_j(1), x_j(2)) / F_{t_{j-1}}(x_j(1), x_j(2))$$

from which the result follows immediately.

We shall now make the assumption that  $G_t$  is differentiable with respect to  $(x_1, x_2, \dots, x_n)$  and  $t$ . Let  $g_t = \partial G_t / \partial t$ .

Theorem 1: For  $0 < s < t$  the multidimensional inhomogeneous extremal process is a step function with only a finite number of jumps in  $[s, t]$ .

Proof: This result is an extension of theorem 4.1 of [2].

Consider  $\{Z_i(t)\}$  and let  $EZ_i(t) = v_i(t)$ . By applying the time transformation

$\tau = f_i^{-1}(t) = e^{v_i(t)}$  the process  $\{V_i(\tau)\} = \{Z_i(f_i(\tau))\}$  becomes a homogeneous extremal process. This is demonstrated in [1], Lemma 3, p. 33. By theorem 4.1 of [2] it follows that  $\{V_i(\tau)\}$  has a finite number of jumps.

To complete the description of the inhomogeneous case we state the transition probabilities and the holding time distribution for a bivariate process,  $\{Z_1(t), Z_2(t)\}$ .

By (2.3) we realize that, given  $Z(s) = (x, y)$ , there will be no jump in  $(s, t)$  if  $\tilde{W}(s, t) < (x, y)$ . But this event has probability  $F_t(x, y)/F_s(x, y)$ . If  $T_s(x, y)$  denotes the holding time in  $(x, y)$ , we therefore have

$$P\{T_s(x, y) > t\} = F_t(x, y)/F_s(x, y).$$

The transition probability function

$$K_{st}(x_2, y_2 | x_1, y_1) \equiv P\{Z_1(t) \leq x_2, Z_2(t) \leq y_2 | Z_1(s) = x_1, Z_2(s) = y_1\}$$

is given by

$$K_{st}(x_2, y_2 | x_1, y_1) = \begin{cases} F_t(x_2, y_2)/F_s(x_1, y_1), & x_2 \geq x_1, y_2 \geq y_1 \\ 0 & \text{otherwise} \end{cases}$$

### 3. The process $\{J(t)\}$

Theorem 2: Let  $\{Z(t)\}$  be a multidimensional inhomogeneous extremal process with marginal distribution  $F_t$  at time  $t$  that satisfies (2.1).

Define the finite state space process  $\{J(t)\}$  as follows:  $J(t) = j$  if  $Z_j(t) = \max_k Z_k(t)$ .

Then the process  $\{J(t)\}$  is an inhomogeneous Markov chain with transition probabilities

$$(3.1) \quad P_{ij}(s, t) \equiv P(J(t) = j | J(s) = i) = - \frac{\partial_j G_t(0) - \partial_j G_s(0)}{G_t(0)}, \text{ for } i \neq j,$$

$$(3.2) \quad P_{ii}(s, t) = \frac{-\partial_i G_t(0) + \partial_i G_s(0) + G_s(0)}{G_t(0)}$$

and state probabilities

$$(3.3) \quad P_j(t) \equiv P(J(t)=j) = -\frac{\partial_j G_t(0)}{G_t(0)}$$

where  $G_t \equiv -\log F_t$  and  $\partial_j$  denotes the partial derivative with respect to the  $j$ -th component.

Remark. Note that when  $i \neq j$  the transition probabilities do not depend on  $i$ . It is in fact this property that allows the aggregation property of Corollary 3.

Proof: Consider first the bivariate case  $\{Z(t)\} = \{Z_1(t), Z_2(t)\}$ .

Let  $0 < t_1 < t_2 < \dots < t_n$  be  $n$  (arbitrary) points in time and let  $\{t_{i_r}\}$  be a subsequence of  $\{t_i, 1 \leq i \leq n\}$ . Put  $E_{i_r} = (t_{i_r-1}, t_{i_r})$  and  $i_k = (i_1, i_2, \dots, i_k)$  where  $i_p \leq i_q$  for  $p \leq q$  and  $i_1 > 1$ . The basic idea of the proof is to consider the probability  $Q_{kn}(i_k)$  defined by

$$Q_{kn}(i_k) = P\left\{ \bigcap_{j=1}^n (Z_1(t_j) \leq Z_2(t_j)) \cap (Z_2(t) \text{ jumps solely in } E_{i_2}, E_{i_3}, \dots, E_{i_k}) \right\},$$

for  $k > 1$ , and

$$Q_{1n} = P\left\{ \bigcap_{j=1}^n (Z_1(t_j) \leq Z_2(t_j)) \cap (Z_2(t_j) = Z_2(t_1), j=2, 3, \dots, n) \right\}.$$

When this probability has been computed for all possible subsequences  $\{t_{i_r}\}$  of  $\{t_i, 1 \leq i \leq n\}$  it is easy to obtain the likelihood of  $\{J(t_1), J(t_2), \dots, J(t_n)\}$ .

Before we start the computation of  $Q_{kn}(i_k)$  we need the following equations

$$(3.4) \quad P\{Z_1(t_j) \leq y, Z_2(t_j) \in dy \mid Z_1(t_k) \leq x, Z_2(t_k) = x, \quad \forall k < j\} \\ = -\exp\{-G_{t_j}(y, y) + G_{t_{j-1}}(y, y)\}(\partial_2 G_{t_j}(y, y) - \partial_2 G_{t_{j-1}}(y, y))dy, \quad x < y$$

and

$$(3.5) \quad P\{Z_1(t_j) \leq y, Z_2(t_j) = y \mid Z_1(t_k) \leq y, Z_2(t_k) = y, \quad \forall k < j\} \\ = \exp\{-G_{t_j}(y, y) + G_{t_{j-1}}(y, y)\}.$$

Eq. (3.4) follows directly from (2.5) and (2.4). Recall that

$Z_2(t_j) \geq Z_2(t_{j-1})$ . Therefore,  $Z_2(t_j)$  cannot be less than  $y$  given that  $Z_2(t_{j-1}) = y$ . This means that  $\{Z_2(t_j) = y\}$  can be replaced by  $\{Z_2(t_j) \leq y\}$  in (3.5) (and vice versa) without altering the probability. Eq. (3.5) now follows from (2.5) and (2.4).

Consider  $Q_{1n}$ . This is the probability that  $\{Z_2(t)\}$  does not jump in  $[t_1, t_n]$  and that  $Z_1(t_j) \leq Z_2(t_j)$  for  $j=1, 2, \dots, n$ . By definition and the fact that  $\{Z_2(t)\}$  is nondecreasing it is clear that we may write

$$Q_{1n} = \int P\left\{\bigcap_{j=2}^n (Z_1(t_j) \leq y, Z_2(t_j) \leq y), Z_1(t_1) \leq y, Z_2(t_2) \in dy\right\}.$$

Decomposing the integrand into conditional probabilities and applying (3.5) give



$$\begin{aligned}
Q_{1n} &= \int \prod_{j=2}^n P\{Z_1(t_j) \leq y, Z_2(t_j) = y | Z_1(t_k) \leq y, Z_2(t_k) = y, \forall k < j\} \\
&\quad \cdot P\{Z_1(t_1) \leq y, Z_2(t_1) \in dy\} \\
&= \int -\exp\left\{-\sum_{j=2}^n (G_{t_j}(y, y) - G_{t_{j-1}}(y, y))\right\} e^{-G_{t_1}(y, y)} \partial_2 G_{t_1}(y, y) dy \\
&= \int -\exp\{-G_{t_n}(y, y)\} \partial_2 G_{t_1}(y, y) dy \\
&= \int -\exp\{-e^{-y} G_{t_n}\} \partial_2 G_{t_1} e^{-y} dy = -\frac{\partial_2 G_{t_1}}{G_{t_n}}
\end{aligned}$$

where we for notational convenience write  $G_t$  instead of  $G_t(0,0)$ . In the last step we have used (2.1) to obtain  $G_t(y, y) = e^{-y} G_t$  and  $\partial_2 G_t(y, y) = e^{-y} \partial_2 G_t$ .

Consider next  $Q_{2n}(i)$ , which is the probability that  $\{Z_2(t)\}$  only jumps in  $(t_{i-1}, t_i)$  and that  $Z_1(t_j) \leq Z_2(t_j)$  for  $j=1, 2, \dots, n$ .

The probability that  $Z_1(t_j) \leq Z_2(t_j)$  for  $j=2, 3, \dots, n$ , and that  $\{Z_2(t)\}$  jumps only in  $(t_{i-1}, t_i)$  from  $x$  into  $(y, y+dy)$  given that  $Z_1(t_1) \leq Z_2(t_1) = x$ , is

$$\begin{aligned}
&P\left\{\bigcap_{j=1}^n (Z_1(t_j) \leq Z_2(t_j)) \bigcap_{r=i}^n (Z_2(t_r) \in dy) \bigcap_{k=2}^{i-1} (Z_2(t_k) = x) \right. \\
&\quad \left. | Z_1(t_1) \leq x, Z_2(t_1) = x\right\}
\end{aligned}$$

which by decomposition into conditional probabilities and application of (3.4) and (3.5) give

$$\begin{aligned}
Q_{2n}(i) &= \int_{x \leq y} \prod_{j=i+1}^n P\{Z_1(t_j) \leq Z_2(t_j) = y | Z_1(t_r) \leq Z_2(t_r) = y, \forall r, i \leq r < j, \\
&\quad Z_1(t_k) \leq Z_2(t_k) = x, \forall k, 1 \leq k < i\} \prod_{j=2}^{i-1} P\{Z_1(t_j) \leq Z_2(t_j) = x | \\
&\quad Z_1(t_k) \leq Z_2(t_k) = x, \forall k, 1 \leq k < i-1\} P\{Z_1(t_1) \leq Z_2(t_1) \in dx\} \\
&\quad \cdot P\{Z_1(t_i) \leq Z_2(t_i) \in dy | Z_1(t_r) \leq Z_2(t_r) = x, \forall r, 1 \leq r < i\} \\
&= \int_{x \leq y} -\exp\left\{-\sum_{j=i+1}^n (G_{t_j}(y, y) - G_{t_{j-1}}(y, y))\right\} \\
&\quad \cdot \exp\{-G_{t_i}(y, y) + G_{t_{i-1}}(y, y)\} (\partial_2 G_{t_i}(y, y) - \partial_2 G_{t_{i-1}}(y, y)) dy \\
&\quad \cdot \exp\left\{-\sum_{j=2}^{i-1} (G_{t_j}(x, x) - G_{t_{j-1}}(x, x))\right\} \cdot \exp\{-G_{t_1}(x, x)\} \partial_2 G_{t_1}(x, x) dx \\
&= \int_{x \leq y} -\exp\{-G_{t_n}(y, y) + G_{t_{i-1}}(y, y)\} (\partial_2 G_{t_i}(y, y) - \partial_2 G_{t_{i-1}}(y, y)) dy \\
&\quad \cdot \exp\{-G_{t_{i-1}}(x, x)\} \partial_2 G_{t_1}(x, x) dx \\
&= \int_{x \leq y} -\exp\{-e^{-y}(G_{t_n} - G_{t_{i-1}})\} (\partial_2 G_{t_i} - \partial_2 G_{t_{i-1}}) e^{-y} dy \\
&\quad \cdot \exp\{-e^{-x} G_{t_{i-1}}\} \partial_2 G_{t_1} e^{-x} dx.
\end{aligned}$$

This final expression reduces to

$$Q_{2n}(i) = -\partial_2 G_{t_1} (\partial_2 G_{t_i} - \partial_2 G_{t_{i-1}}) / G_{t_{i-1}} G_{t_n}.$$

Let

$$(3.6) \quad M_i \equiv \frac{\partial_2 G_{t_i} - \partial_2 G_{t_{i-1}}}{G_{t_{i-1}}}, \quad i \geq 2.$$

By the same procedure as above we obtain

$$Q_{2n}(i,j) = \frac{-\partial_2 G_{t_1}^{M_i M_j}}{G_{t_n}}, \quad i < j,$$

and in the general case

$$(3.7) \quad Q_{kn}(i_k) = \frac{-\partial_2 G_{t_1}^k}{G_{t_n}} \prod_{p=1}^k M_{i_p}.$$

Let

$$Q_n = P \left\{ \bigcap_{j=1}^n (Z_1(t_j) \leq Z_2(t_j)) \right\}.$$

Then obviously

$$(3.8) \quad Q_n = \sum_{k=1}^n \sum_{i_k} Q_{kn}(i_k)$$

because

$$\sum_{i_k} Q_{kn}(i_k)$$

is the probability that  $Z_2(t_j) \geq Z_1(t_j)$ ,  $j=1,2,\dots,n$ , and that  $Z_2(t)$  jumps in  $k$  of the intervals  $(t_{j-1}, t_j)$ ,  $j=2,3,\dots,n$ . Now by (3.7) and (3.8) we get

$$(3.9) \quad Q_n = \frac{-\partial_2 G_{t_1}}{G_{t_n}} \left( 1 + \sum_{k=1}^n \sum_{i_k} \prod_{p=1}^k M_{i_p} \right).$$

From classical algebra we have the identity

$$(3.10) \quad 1 + \sum_{k=1}^n \sum_{i_k} \prod_{p=1}^k M_{i_p} = \prod_{j=2}^n (1 + M_j).$$

Moreover, by (3.6)

$$\begin{aligned} \prod_{j=2}^n (1+M_j) &= \prod_{j=2}^n \left( \frac{G_{t_{j-1}} - \partial_2 G_{t_j} + \partial_2 G_{t_{j-1}}}{G_{t_{j-1}}} \right) \\ &= \frac{G_{t_n}}{G_{t_1}} \prod_{j=2}^n \frac{G_{t_{j-1}} - \partial_2 G_{t_j} + \partial_2 G_{t_{j-1}}}{G_{t_j}} \end{aligned}$$

which by (3.9) and (3.10) implies that

$$(3.11) \quad Q_n = - \frac{\partial_2 G_{t_1}}{G_{t_1}} \prod_{j=2}^n \frac{G_{t_{j-1}} - \partial_2 G_{t_j} + \partial_2 G_{t_{j-1}}}{G_{t_j}}.$$

The probability of  $(Z_1(t) < Z_2(t))$  is found by straightforward integration and application of (2.1) to be

$$(3.12) \quad P\{Z_1(t) < Z_2(t)\} = \int -\exp\{-e^{-y} G_t\} e^{-y} \partial_2 G_t dy = \frac{-\partial_2 G_t}{G_t}$$

which proves (3.3).

A consequence of (3.10) and (3.12) is that

$$\begin{aligned} &P\{Z_1(t_n) \leq Z_2(t_n) | Z_1(t_j) \leq Z_2(t_j), \quad j=1,2,\dots,n-1\} \\ &= P\{Z_1(t_n) \leq Z_2(t_n) | Z_1(t_{n-1}) \leq Z_2(t_{n-1})\}. \end{aligned}$$

Since this is true for any  $\{t_j, j \leq n\}$  it implies that  $\{J(t)\}$  is a Markov chain. From (3.10) we also get

$$P\{Z_1(t_j) \leq Z_2(t_j) | Z_1(t_{j-1}) \leq Z_2(t_{j-1})\} = \frac{G_{t_{j-1}} - \partial_2 G_{t_j} + \partial_2 G_{t_{j-1}}}{G_{t_j}}$$

which yields (3.2) and (3.1). Hence, the theorem is proved in the

bivariate case when  $J(t_j)=1$  or  $J(t_j)=2$  for  $j=1,2,\dots,n$ . But then the theorem must also hold in the general bivariate case because the likelihood of a general sample path can be expressed by joint probabilities of being in state  $j$ , ( $j=1,2$ ) at some points in time. For instance,

$$\begin{aligned} & P\{J(t_1)=1, J(t_2)=1, J(t_3)=2\} \\ &= P\{J(t_1)=1, J(t_2)=1\} - P\{J(t_1)=J(t_2)=J(t_3)=1\}. \end{aligned}$$

Now it is easily verified that the transition probabilities of  $\{J(t)\}$  satisfy the Chapman - Kolmogorov equations. Hence, there exists a Markov chain defined by these transition probabilities. Since the transition probabilities uniquely characterize a Markov process and the likelihood  $Q_n$  can be expressed by the transition probabilities, the likelihood in the general (bivariate) case must also satisfy the Markov property.

In the general case where the dimension of  $\{Z(t)\}$  is greater than two the theorem is proved in the same way as in [1], p.p.

41-42. The essential property used in the rest of the proof is that

$\{Z_i(t), \max_{k \neq i} Z_k(t)\}$  is also a bivariate extremal process. This property follows directly from assumption (2.1).

This completes the proof.

Theorem 2 tells us that we can define a discrete state space Markov chain  $\{J(t)\}$  from the continuous state Markov process  $\{Z(t)\}$  where the transition probabilities are given by (3.1) and (3.2).

Corollary 1: The transition probabilities of the Markov chain  $\{J(t)\}$  can be expressed as

$$P_{ij}(s,t) = P_j(t) - P_j(s)\zeta(s,t), \quad i \neq j$$

where

$$\zeta(s,t) \equiv \text{corr}\{\exp(-\max_k Z_k(t)), \exp(-\max_k Z_k(s))\} = \frac{G_s(0)}{G_t(0)}.$$

Proof: By (3.1) and (3.3) we have

$$P_{ij}(s,t) = P_j(t) - P_j(s) \frac{G_s(0)}{G_t(0)}$$

which proves the first part of the corollary.

Since  $\{\max_k Z_k(t)\}$  is a univariate extremal process, it follows that

$$\{\exp\{-\max_k Z_k(s)\}, \exp\{-\max_k Z_k(t)\}\}$$

is bivariate exponentially distributed. From [5] we get

that the autocorrelation function of  $\exp\{-\max_k Z_k(t)\}$  at  $(s,t)$  is  $G_s(0) / G_t(0)$ .

This completes the proof.

The interest of Corollary 1 is that it expresses the transition probabilities in terms of the state probabilities and a term,  $\zeta(s,t)$ , that is a measure of the temporal stability of  $\{\max_k Z_k(t)\}$ .

The next corollary concerns the transition intensities of  $\{J(t)\}$ . Recall that the transition intensities are defined by

$$\lambda_{ij}(t) = \lim_{\Delta t \rightarrow 0} \frac{P_{ij}(t, t+\Delta t) - P_j(t)}{\Delta t} \quad \text{for } i \neq j$$

and

$$\lambda_{ii}(t) = \lim_{\Delta t \rightarrow 0} \frac{P_{ii}(t, t+\Delta t) - 1}{\Delta t}.$$

Corollary 2: The Markov chain  $\{J(t)\}$  has transition intensities

$$\lambda_{ij}(t) = \frac{-\partial_j g_t(0)}{G_t(0)} \quad \text{for } i \neq j$$

and

$$\lambda_{ii}(t) = -\sum_{k \neq i} \lambda_{ik}(t) = \frac{-\partial_i g_t(0) - g_t(0)}{G_t(0)}.$$

The excursion time of  $Z_i(t) - \max_{k \neq i} Z_k(t)$  has distribution

$$\begin{aligned} & P\{ \inf_{s \leq \tau \leq t} (Z_i(\tau) - \max_{k \neq i} Z_k(\tau)) > 0 \mid J(s)=i \} \\ &= \exp\left\{ \int_s^t \lambda_{ii}(x) dx \right\} = \frac{G_s(0)}{G_t(0)} \exp\left\{ -\int_s^t \frac{\partial_i g_\tau(0)}{G_\tau(0)} d\tau \right\}. \end{aligned}$$

Proof: By Theorem 2 we get  $\lambda_{ij}(t)$  for  $i \neq j$  by differentiation.

Notice that since

$$1 = \sum_j P_j(t) = -\sum_j \frac{\partial_j G_t(0)}{G_t(0)},$$

we have

$$-\sum_j \partial_j G_t(0) = G_t(0).$$

By using this result we get the expression for  $\lambda_{ii}(t)$ . Since  $\{J(t)\}$  is a Markov chain the last result follows immediately. This completes the proof.

A particular feature of  $\{J(t)\}$  is that its structure is invariant under aggregation of states. This is a consequence of the fact that the class of multidimensional extremal processes is invariant under maximization of components of the process. We state this result below.

Corollary 3: The family of Markov chains  $\{J(t)\}$  is invariant under aggregation of states.

As mentioned above condition (2.1) implies that the distribution of  $Z_j(t)$  is extreme value type III. It is, however, easily realized that Theorem 2 holds for more general distributions of  $Z_j(t)$ . In fact we have

Corollary 4: Let  $\{\tilde{U}(t)\}$  be a multidimensional extremal process with general one-dimensional marginal distributions. Let  $F_t$  be the distribution of  $\tilde{U}(t)$  and  $G_t = -\log F_t$ . Assume that there exists a family of increasing functions  $\{\Psi_t(x), t \geq 0\}$  such that  $\tilde{G}_t$  defined by

$$\tilde{G}_t(x_1, x_2, \dots, x_m) = G_t(\Psi_t(x_1), \Psi_t(x_2), \dots, \Psi_t(x_m))$$

satisfies condition (2.1). Then Theorem 2 holds with  $G_t$  replaced by  $\tilde{G}_t$ .

Proof: Define  $\{\tilde{Z}(t)\}$  by

$$\tilde{Z}(t) = (\Psi_t^{-1}(U_1(t)), \Psi_t^{-1}(U_2(t)), \dots, \Psi_t^{-1}(U_m(t))).$$

Then  $\tilde{Z}(t)$  has distribution  $\exp(-\tilde{G}_t)$ . Now observe that  $\{U_i(t) = \max_k U_k(t)\}$  is equivalent to  $\{Z_i(t) = \max_k Z_k(t)\}$  because  $\Psi_t$  is increasing. Hence, the claims of the corollary follow from Theorem 2 and the proof is complete.

#### Example

Let  $G_t = e^{\theta t} G$  where  $\theta > 0$  is a constant and let  $\{Z^*(t)\}$  be the corresponding process. The one-dimensional version of  $\{Z^*(t)\}$  has been studied by Tiago de Oliveira [7]. The process  $\{Z^*(t) - \theta t\}$  is stationary which is easily verified by checking the corresponding finite dimensional marginal distributions. Tiago de Oliveira calls this process (the one-dimensional version) the extreme Markovian stationary process. Let  $\{J^*(t)\}$  be the (homogeneous) Markov chain generated by  $\{Z^*(t)\}$ .



From Theorem 2 we get the state and the transition probabilities

$$P_j^* = \frac{\partial_j G(0)}{G(0)}$$

and

$$P_{ij}^*(s, t) = P_j^*(1 - e^{-\theta(t-s)}) \quad \text{for } i \neq j.$$

From Corollary 2 we get the holding time distribution of state  $i$ :

$$\begin{aligned} & P\left\{ \inf_{s \leq \tau \leq t} (Z_i^*(\tau) - \max_{k \neq i} Z_k^*(\tau)) > 0 \mid J^*(s) = i \right\} \\ &= \exp\{- (t-s)\theta(1-P_i^*)\}. \end{aligned}$$

When the components  $Z_1^*(t), Z_2^*(t), \dots$  are independent, then

$$G(\underline{x}) = \sum_k e^{v_k - x_k}$$

where  $v_k = EZ_k^*(t) - \theta t$ . Hence we get

$$P_j^* = \frac{e^{v_j}}{\sum_k e^{v_k}}.$$

Thus in this case the state probabilities are multinomial logit functions of the parameters  $v_k$ .

#### 4. Applications

The results derived above are, as mentioned, of particular interest for applications in economics and psychology because they provide a framework for analyzing the structure of individual discrete decisions over time.

Consider the analysis of individual migration careers. Let  $Z_j(t)$  be the individuals' utility of being in region  $j$  at age  $t$ . The individual decision rule is to stay in the region with the highest utility. Thus a move takes place each time the utility of another region becomes higher than the utility of the region in which the individual stays for the moment. The utility  $Z_j(t)$  may be a function of individual characteristics as well as characteristics of region  $j$ , for instance, employment rate, urbanization, etc. Since only some of the variables that influence the choice process are observable to the observer, the utility function is random. Also the utility function may be correlated over time because of temporal stability in unobserved factors.

If the utilities are assumed to be extremal processes, the above results enable us to express the transition intensities of the observed migration process as functions of the parameters of the individuals' utility processes. The choice of the extremal process can also be given a behavioral justification (cf. [1]).

The above model framework can be used to discriminate between two different explanations for observed dependence on previous migration states. One is called "true state dependence" and the other is called "habit persistence" or "heterogeneity".

The first explanation, "true state dependence", is that past experience has a genuine behavioral effect in the sense that the behavior of otherwise identical individuals who did not have the same experience would be different in the future. The other explanation, heterogeneity, is that individuals may differ in their propensity to experience certain careers. If individual differences are correlated over time and if these differences are not properly controlled, previous experience may appear to be a determinant of future experience solely

because it is a proxy for temporally persistent unobservables that determine choices.

In the example at the end of section 3 the heterogeneity or habit persistent effect is represented by the parameter  $\theta$ . If  $\theta$  is large the temporal stability in the unobservables is weak while when  $\theta$  is small the "habit persistence" is strong. The state dependence effects may be modelled through expected utilities by letting  $v_j$  depend on previous realizations of the migration process.

For a more detailed discussion of these modelling issues the reader is referred to [4].

## References

- [1] J. Dagsvik, Discrete dynamic choice. An extension of the choice models of Thurstone and Luce. J. Math. Psychology, 27 (1983), 1-43.
- [2] M. Dwass, Extremal processes. Ann. Math. Statist, 35 (1964), 1718-1725.
- [3] L. de Haan and S.I. Resnick, Limit theory for multivariate sample extremes. Z. Wahrscheinlichkeit. Verw. Gebiete, 40 (1977), 317-337.
- [4] J.J. Heckman, Statistical models for the analysis of discrete panel data. In C.F. Manski and D. McFadden (Eds.) Structural analysis of discrete data. Cambridge, MIT Press (1981).
- [5] N. Johnson, and S. Kotz, Distribution in statistics: Continuous multivariate distributions. New York, Wiley (1972).
- [6] R.D. Luce, Individual choice behavior. New York, Wiley (1959).
- [7] D. McFadden, Probabilistic theories of choice. In C.F. Manski and D. McFadden (Eds.) Structural analysis of discrete data. Cambridge, MIT Press (1981).
- [8] J. Tiago de Oliveira, An extreme Markovian stationary process. Proceedings of the fourth conference in probability theory, Acad. Romania, Brasov, (1973) 217-225.
- [9] I. Weissman, Extremal processes generated by independent non-identically distributed random variables. Ann. Probability, 3 (1975), 172-177.
- [10] I. Weissman, Multivariate extremal processes generated by independent nonidentically distributed random variables. J. Appl. Probability, 12 (1975), 447-487.

ISSUED IN THE SERIES DISCUSSION PAPER

- No. 1 I. Aslaksen and O. Bjerkholt: Certainty Equivalence Procedures in the Macroeconomic Planning of an Oil Economy.
- No. 2 E. Offerdal: Structural Change, Profitability, and Survival.
- No. 3 E. Biørn: On the Prediction of Population Totals from Sample Surveys Based on Rotating Panels.
- No. 4 P. Frenger: A Short Run Dynamic Equilibrium Model of the Norwegian Production Sectors.
- No. 5 I. Aslaksen and O. Bjerkholt: Certainty Equivalence Procedures in Decision-Making under Uncertainty: an Empirical Application.
- No. 6 E. Biørn: Depreciation Profiles and the User Cost of Capital.
- No. 7 P. Frenger: A Directional Shadow Elasticity of Substitution.
- No. 8 S. Longva, L. Lorentsen, and Ø. Olsen: The Multi-Sectoral Model MSG-4, Formal Structure and Empirical Characteristics.
- No. 9 J. Fagerberg and G. Sollie: The Method of Constant Market Shares Revisited.
- No. 10 E. Biørn: Specification of Consumer Demand Models with Stochastic Elements in the Utility Function and the First Order Conditions.
- No. 11 E. Biørn, E. Holmøy, and Ø. Olsen: Gross and Net Capital, Productivity, and the Form of the Survival Function - Some Norwegian Evidence.
- No. 12 J. K. Dagsvik: Markov Chains Generated by Maximizing Components of Multidimensional Extremal Processes.